

AperTO - Archivio Istituzionale Open Access dell'Università di Torino

Tullio De Mauro e la lessicografia

This is a pre print version of the following article:

Original Citation:

Availability:

This version is available <http://hdl.handle.net/2318/1671238> since 2018-09-22T21:50:40Z

Publisher:

il Mulino

Terms of use:

Open Access

Anyone can freely access the full text of works made available as "Open Access". Works made available under a Creative Commons license can be used according to the terms and conditions of said license. Use of all other works requires consent of the right holder (author or publisher) if not exempted from copyright protection by the applicable law.

(Article begins on next page)

CARLA MARELLO

TULLIO DE MAURO E LA LESSICOGRAFIA

1. *Introduzione*

Molto prima che vedessero la luce, nell'arco di cinque anni dal 1996 al 2000, il *Dizionario di base della lingua italiana* (DIB), il *Dizionario avanzato dell'italiano corrente* (DAIC), il *Grande dizionario italiano dell'uso* in sei volumi nel 1999 presso la Utet e il *Dizionario della lingua italiana*, in un volume¹ presso Paravia, De Mauro non solo aveva già guidato imprese di natura lessicologica e lessicografica, ma aveva partecipato alla stesura di voci di dizionario, come ci ha raccontato nell'articolo *Memorie del GRADIT*²:

Del *Dizionario enciclopedico italiano* della Treccani avevo avuto modo di conoscere dall'interno i pregi, perché dai primi anni Sessanta per una dozzina d'anni avevo lavorato, come redattore per la linguistica, le trascrizioni fonetiche e le etimologie, a quella che doveva essere una seconda edizione e poi, per ragioni commerciali, diventò un'opera autonoma, il *Lessico universale italiano* [De Mauro 2017, 8].

In questo saggio non possiamo certo trattare estesamente di tutti i suoi dizionari, ma ci soffermeremo maggiormente

¹ Volume unico, ma di ben 2.998 pagine più XXV di introduzione: una specie di record nel mercato italiano e non solo, se si considera che l'ultima edizione cartacea del dizionario Zingarelli 2018 raggiunge 2688 pagine più poche altre non numerate. Nel 2004 Paravia commercializzò una versione più maneggevole, *Il dizionario di italiano compatto*: 65.000 voci e 1700 pagine.

² Pubblicato nella Strenna Utet a cura di Claudio Marazzini [De Mauro 2017] (e subito ristampato, per consentirne una più ampia diffusione, nella rivista «Lingua e stile») è uno scritto a cui ci rifaremo spesso, perché va ben al di là delle vicende di composizione del *Grande dizionario italiano dell'uso*, abbreviato in GRADIT, costituisce anzi una sorta di autobiografia di De Mauro in quanto lessicografia.

sugli aspetti della sua lessicografia e metalessicografia che più dovrebbero far parte della formazione degli insegnanti di italiano e sulle opere che, essendo in rete, con accesso libero, possono raggiungere una gran parte di italiani.

Abbiamo menzionato la metalessicografia di De Mauro perché la riflessione sui dizionari, e in particolare sui dizionari che costruiva, costituisce una cospicua parte della sua vasta bibliografia: il paratesto del GRADIT tra pre- e postfazioni³ ha raggiunto le dimensioni di volumi a sé e uno dei suoi ultimi scritti, datato 23 dicembre 2016 ore, 13.01, è la presentazione del *Nuovo vocabolario di base della lingua italiana*⁴ (d'ora in poi NVdB), testo tanto sociolinguistico, quanto metalessicografico.

Se le analisi quantitative dei fatti di lingua l'hanno sempre appassionato, va però sottolineato che la sua era una linguistica quantitativa strettamente al servizio della sociolinguistica e della linguistica teorica. Nei suoi scritti sui dizionari (e in genere nei suoi scritti: valga per tutti il titolo *Parole e numeri. Analisi quantitative dei fatti di lingua*)⁵ ha sempre cercato di spiegare perché è importante servirsi dei dati quantitativi e perché è ancora più importante riuscire a fare una lessicografia che partendo dai numeri restituisca l'uso della massa parlante. I dizionari dovevano giustamente apparirgli come il modo più capillare e «normale», meno appariscente e quindi più efficace⁶, di instillare una visione

³ Un aspetto che Rovere [2003, 2] sottolinea in quanto dettagliata esposizione del modo di intendere il dizionario e dei correlati problemi teorici e metodologici, ben diversa dalle solite prefazioni.

⁴ Si può leggere per ora in rete nella rivista «Internazionale», <https://www.internazionale.it/opinione/tullio-de-mauro/2016/12/23/il-nuovo-vocabolario-di-base-della-lingua-italiana> consultato il 30 agosto 2017.

⁵ A cura di Isabella Chiari e Tullio De Mauro [2005].

⁶ Non si evoca l'immagine del cavallo di Troia proprio perché appariscente; tale immagine è però opportunamente usata nella chiusa del già citato *Memorie del GRADIT* da De Mauro stesso, a conferma dell'importanza che attribuiva al tirocinio (meta)lessicografico. Dice infatti che dal lavoro redazionale del GRADIT e delle imprese precedenti sono usciti «tamquam equus Troianus» linguisti che in seguito hanno dato contributi significativi agli studi linguistici in Italia, quali, fra gli altri, Federica Casadei, Claudio Iacobini, Luca Lorenzetti, Marco Mancini,

(socio)linguistica fra gli utenti, soprattutto fra coloro che non hanno avuto, né avranno la ventura di frequentare corsi di linguistica.

I linguisti e i metalessicografi più impegnati potrebbero rimproverargli di esser stato troppo cauto, di aver accettato una descrizione per parte del discorso molto tradizionale⁷, ma nelle scelte che ha fatto⁸ si nota che il suo interesse era concentrato sull'uso, in quanto traspariva da e incideva su semantica e pragmatica del lessico italiano. Il *Vocabolario di base* (VdB) costituì un segnale importante: dopo la sua pubblicazione gli altri dizionari cominciarono a marcare i lemmi relativi alle parole incluse nel VdB⁹ e nel demauriano GRADIT le parole del VdB sono state scelte «per costruire l'inizio delle definizioni del significato dei lemmi in modo che almeno l'incipit delle definizioni risulti chiaro a un largo insieme di lettori e lettrici, lasciando al seguito della definizione specificazioni che possano richiedere parole di uso più raro e di minore circolazione»¹⁰.

De Mauro fece entrare nel panorama italiano quell'at-

Raffaella Petrilli, Emanuela Piemontese, Massimo Vedovelli (citati, si badi, in ordine rigorosamente alfabetico).

⁷ Pare di leggere questa osservazione in Rovere [2003, 14] quando osserva che il GRADIT si presenta come il luogo della descrizione generale della lingua italiana, almeno come obiettivo finale, ma non ha accolto l'innovazione dei criteri valenziali introdotti in lessicografia dal Sabatini-Coletti [1997], né aggiungiamo le costruzioni di Peter Blumenthal e Giovanni Rovere, *Wörterbuch der italienischen Verben* (Dizionario dei verbi italiani) appena uscito nel 1998.

⁸ Tenendo anche conto delle prevedibili cautele richieste dalle case editrici, le innovazioni demauriane più appariscenti nel metalinguaggio sono state *polirematica* e *procomplementare*. Le polirematiche sono espressioni che constano di più parole e sono dotate di un significato unitario non ricavabile per semplice somma dei significati degli elementi costitutivi, per es. *bagaglio a mano*, *andarci di mezzo*. Procomplementare è detto un elemento pronominale che sostituisce un gruppo nominale con funzione di complemento; un verbo procomplementare è un verbo che si usa stabilmente con particelle clitiche procomplementari, ad es. *svignarsela*, o un verbo che può usarsi con tali particelle, ad es. *andar-sene*, e che può assumere significati autonomi rispetto al verbo di base, ad es. *sentirsela*, *vedersela*.

⁹ Di solito allargandone il numero dalle originali 7.050 fino a 10.000.

¹⁰ Citiamo sempre dalla presentazione on line del NVdB.

tenzione per il lessico più comune che nel frattempo i dizionari per apprendenti di inglese e francese, con i loro lemmari ben calcolati e le loro definizioni costruite con le parole di lessico fondamentale, avevano diffuso fra i docenti di lingue straniere.

2. *Il Vocabolario di base*

È importante notare che De Mauro ha sempre usato il termine *vocabolario*¹¹ per le opere che consistevano di liste di lemmi con al massimo l'indicazione della parte del discorso, ma senza sfociare in voci con definizioni ed esempi. La lista del *Vocabolario di base* riportata nel libro *Guida all'uso delle parole* [De Mauro 1980, 149-183] è il risultato di ricerche condotte da Stefano Gensini e Emilia Passaponti a partire dal mondo della scuola, nei suoi vari gradi, e collegato agli studi della carta d'identità linguistica felicemente battezzata «glotto-kit».

All'epoca quando circolavano i questionari che chiedevano agli intervistati di indicare se conoscevano o non conoscevano una parola, se non l'avevano mai sentita o se l'avevano sentita ma non la conoscevano, oppure se la conoscevano ma non la usavano e infine se la conoscevano e usavano, c'era un certo scetticismo sulla sincerità delle risposte a domande che, tra l'altro, sapevano di «esame di coscienza», sia pure linguistica. E invece gli intervistati furono onesti e/o gli intervistatori abbastanza abili da neutralizzare con le tecniche psicometriche e statistiche del caso le eventuali risposte insincere date per apparire più forbiti e colti. Il VdB si rivelò uno strumento potente per tutti coloro che scrivevano libri di testo scolastici o semplicemente libri che volevano essere leggibili per una grande parte della popolazione¹².

¹¹ O *lessico* come nel caso del *Lessico dell'italiano parlato* (LIP). Quando invece definisce le parole, anche se non sono molte, usa *dizionario*.

¹² Ricorda De Mauro che: «case editrici, come la Giunti, si sono servite del VdB per orientare al meglio la redazione e comprensibilità

Il sociolinguista De Mauro meglio di chiunque altro sapeva che l'uso cambia, che il vocabolario di ogni individuo muta con il passare degli anni e perciò si rendeva necessario un *Nuovo vocabolario di base* che ci ha regalato a Natale del 2016, gratuitamente scaricabile da un link all'inizio della presentazione¹³.

Il NVdB si fonda sullo spoglio elettronico (controllato manualmente) di quasi 19 milioni di occorrenze in testi di sei tipi: stampa (quotidiani e settimanali), saggistica (saggi divulgativi, testi e manuali scolastici e universitari), testi letterari (narrativa, poesia), spettacolo (copioni cinematografici e televisivi, teatro), comunicazione mediata dal computer (chat eccetera), registrazioni di parlato¹⁴.

Oggi troviamo ovvio che i libri di scuola siano scritti con un lessico controllato, ma come si giunga a tale lessico controllato sfugge ai più: al massimo si pensa alla frequenza, che certo è importante, ma non può essere il solo criterio per chi si pone obiettivi di formazione.

2.1. VELI e LIP

La feconda collaborazione fra Ibm e umanisti avviata da Padre Busa nel lontano 1949 era destinata con De Mauro a produrre altri frutti italiani negli anni Novanta del secolo scorso. Il *Vocabolario elettronico della lingua italiana* (VELI, 1989), fu realizzato dalla Direzione Ricerca Scientifica e Tecnologica Ibm Italia sotto la guida di De Mauro: è costituito

di testi per la scuola primaria. Dal 1992, su iniziativa del ministro della funzione pubblica Sabino Cassese, il VdB fu assunto a riferimento del *Codice di stile delle comunicazioni dell'amministrazione pubblica*. Tra il 1980 e il 1989 gli autori della collana «Libri di base» degli Editori Riuniti se ne servirono anche per introdurre «in modo comprensibile le molte parole estranee al VdB e tuttavia utili alla trattazione dei temi più diversi» (dalla presentazione on line del NVdB).

¹³ <https://www.dropbox.com/s/mkcyo53m15ktbnp/nuovovocabolarioibase.pdf?dl=0>, consultato il 30 agosto 2017.

¹⁴ Informazioni interessanti sul NVdB si trovano in Chiari e De Mauro [2012; 2014].

da circa 10.000 lessemi ordinati per frequenza nella lingua italiana. La frequenza è calcolata a partire da un corpus di più di 25 milioni di parole, quindi enormemente più esteso del corpus alla base del LIF (che era di sole 500.000 parole)¹⁵, ma meno bilanciato, dal momento che le 10.000 parole sono tratte da fonti giornalistiche (Ansa, «Il Mondo», «Europeo», «Domenica del Corriere») nel periodo 1985-1987.

Come osserva Isabella Chiari nel suo sito¹⁶, il VELI «già dispone di un insieme di strumenti di trattamento dei dati testuali di seconda generazione [...] tuttavia l'accesso al corpus non è dato all'utente». Con il LIP, *Lessico di frequenza dell'italiano parlato*, curato da De Mauro, Mancini, Vedovelli e Voghera e pubblicato nel 1993, l'utente ha invece avuto l'accesso¹⁷ al corpus di trascrizioni di registrazioni effettuate a Milano, Firenze, Roma e Napoli. Circa 57 ore di parlato – monologhi, dialoghi faccia a faccia e telefonici – che confluiscono, trascritte, in 500.000 parole grafiche, la controparte parlata del LIF. Grazie al sito della Karl-Franzens-Universität di Graz (Austria)¹⁸ è ora disponibile gratuitamente una versione online del corpus del *Lessico di frequenza dell'italiano parlato*. Disponiamo anche del VoLIP, «Voce del LIP» con file audio e una revisione di alcuni dei problemi di trascrizione della versione originale¹⁹: permette la ricerca nel corpus sia secondo criteri sociolinguistici sia secondo criteri lessicali e morfo-sintattici.

A noi che oggi dettiamo a cellulari in grado di «capirci» sempre meglio, grazie a perfezionati programmi di riconoscimento vocale, può esser utile sapere quanto avveniva

¹⁵ Il *Lessico di frequenza della lingua italiana contemporanea* fu tratto da testi di teatro, romanzi, copioni cinematografici, periodici, sussidari. Quando il LIF fu compilato, nel 1970, non si era ancora diffusa la composizione digitale dei testi da stampare, procedimento che ha reso in seguito molto meno dispendioso procurarsi e preparare i testi da sottoporre ai software di trattamento delle lingue naturali.

¹⁶ www.alphabit.net, consultato il 30 agosto 2017.

¹⁷ Attraverso i due dischetti che corredevano il volume a stampa contenente le liste, le frequenze e vari studi.

¹⁸ <http://badip.uni-graz.at/it>, consultato il 30 agosto 2017.

¹⁹ Si veda <http://www.parlaritaliano.it/index.php/it/volip>, consultato il 30 agosto 2017, e Voghera *et al.* [2014].

25 anni fa. Lasciamo la parola a Massimo Vedovelli che, da testimone oculare, ha avuto la cortesia di raccontarci la genesi del LIP.

Ero presente al colloquio che i vertici del centro ricerche Ibm Italia ebbero con De Mauro. Gli rappresentavano questa situazione: i loro gruppi di ricerca avevano come obiettivo di realizzare il dettatore automatico, soprattutto per specifiche categorie professionali (i radiologi, che avrebbero potuto dettare al PC le loro diagnosi guardando le radiografie e non scrivendo), ma anche per una più generalizzata spendibilità. I loro giovani ricercatori applicavano senza successo i modelli (allora disponibili) di intelligenza artificiale. I loro giovani ricercatori il venerdì pomeriggio, prima di lasciare il lavoro, giocavano la schedina del Totocalcio applicando un «sistemino», ovvero un programmino statistico-probabilistico (non so con quale successo, ma spero in un buon successo per quegli allora giovani ricercatori). Un giorno, uno di loro, frustrato dai pessimi risultati dell'intelligenza artificiale, che non riconosceva niente di ciò che veniva dettato al computer, applica il programmino che usavano per il sistemino al Totocalcio. Miracolo: il programma riconosce più parole rispetto a quello basato sui modelli dell'intelligenza artificiale! Insomma, tirando a dadi, il programma dettava meglio al PC!

All'Ibm viene un'idea: mettiamo da parte l'intelligenza artificiale, facciamo tirare a dadi al programma, ma facciamolo scegliere fra un lessico ristretto rispetto a quello generale della lingua italiana, appunto concentriamolo sul lessico del parlato. Ma qual è questo lessico del parlato? Allora, pensano di andarlo a chiedere al Prof. De Mauro, che risponde – io c'era – «e chi lo sa!» Da qui, la proposta di De Mauro di fare il LIP, con ben piccola spesa per la Ibm, che accetta. Il resto, lo conosciamo...

Il LIP ha ravvivato la discussione sul lessico fondamentale e inaugurato una serie di confronti fra le parole più frequenti nel LIP e le parole più frequenti dei precedenti lessici di frequenza sia di italiano scritto sia di parlato²⁰ e, all'interno

²⁰ Cfr. Bortolini, Tagliavini e Zampolli [1972]; Juilland e Traversa [1973]; Reiske [1974]; Baldelli e Mazzetti [1974]; Sciarone [1977]; Direzione Ricerca Scientifica e Tecnologica Ibm Italia [1989]; Katerinov, Boriosi e Sciarone [1991]. Per i confronti con il VdB si veda l'introduzione di Vedovelli [1993] e De Palo [1997].

del gruppo di collaboratori di De Mauro, ha promosso studi sul confronto fra il LIP e i lemmi delle fasce FO, AU, AD²¹ nel *Vocabolario di base*.

3. *Una marca d'uso per ogni accezione: il GRADIT, Il Dizionario della lingua italiana e il Nuovo De Mauro*

Chi scrive ha imparato la parola *gheriglio* a novembre della prima elementare (che allora iniziava a ottobre), perché una lettura dell'abecedario parlava dei frutti dell'autunno, castagne con i loro ricci e noci con mallo, guscio e gheriglio²². *Gheriglio* è parola mai entrata nel *Vocabolario di base* e – l'ho verificato – assente anche dal vocabolario ricettivo di molti universitari; è però parola inserita da De Mauro fra le 15.000 voci del *Dizionario di base*: non è ovviamente fra le 7.050 FO fondamentali, AU molto frequenti, AD alta disponibilità, ma è fra le altre 8.000 che ricevono l'etichetta CO (comune) nel DIB.

De Mauro attribuisce la marca d'uso CO comune alle parole che «sulla scorta dell'esperienza didattica e dell'osservazione risultano indispensabili per la comprensione dell'informazione corrente giornalistica, di costume, sportiva, di testi letterari contemporanei, narrativi e storici, di avvisi al pubblico, moduli, manuali di istruzioni»; parole costituenti «un corredo che permette un dominio affinato e responsabile dei mezzi espressivi e della comprensione, adeguato alle più importanti necessità di vita in una società moderna

²¹ Si ricordano le marche d'uso più importanti attribuite ai vocaboli del VdB: con FO, fondamentale, sono marcati i vocaboli di altissima frequenza, le cui occorrenze costituiscono circa il 90% delle occorrenze lessicali nell'insieme di tutti i testi scritti o discorsi parlati; con AU, alto uso, sono marcati i vocaboli di alta frequenza, le cui occorrenze costituiscono un altro 6% circa delle occorrenze lessicali nell'insieme di tutti i testi scritti o discorsi parlati; con AD, alta disponibilità, sono marcati i vocaboli, relativamente rari nel parlare o scrivere, ma tutti ben noti perché legati ad atti e oggetti di grande rilevanza nella vita quotidiana (*alluce*, *batuffolo*, *carrozzeria*, *dogana*, ecc.).

²² La parola era inserita anche perché contiene le sequenze di lettere *gh* e *gl*, difficili per chi sta imparando a scrivere l'italiano.

e complessa» [De Mauro e Moroni 1996]. Nel *Dizionario della lingua italiana* in un volume (2000) su 130.000 voci ben 40.000 lemmi semplici e 3.500 polirematiche hanno la marca CO (dall'Introduzione, p. IX). Nel *Compatto* del 2004 su 65.000 voci «la selezione e l'alleggerimento hanno inciso progressivamente sugli strati più esterni. Di meno sull'ampia fascia dei vocaboli «comuni», per lo più usati e compresi da chiunque abbia una buona scolarizzazione di livello secondario superiore. Di più sui vocaboli «letterari» e soprattutto sul vastissimo campo dei termini tecnico-specialistici» (dall'introduzione firmata «l'Editore»). Nelle Avvertenze del *Nuovo De Mauro* è detto «comune; sono così marcati i vocaboli che sono usati e compresi indipendentemente dalla professione o mestiere che si esercita o dalla collocazione regionale e che sono generalmente noti a chiunque abbia un livello mediosuperiore di istruzione».

Si comprenderà che CO è un'etichetta più soggetta a discussione²³ di FO, AU, AD, perché molto legata alle variazioni del bagaglio culturale condiviso. Tuttavia anche in questo caso il dibattito ha provocato una sana presa di coscienza su che cosa voglia dire conoscere una parola quando questa ha più di un'accezione, caso più frequente di quanto non si creda, soprattutto con un lemmario di 15.000-20.000 parole scelte anche in base alla frequenza e alla dispersione²⁴.

In quanto lista di lemmi non definiti²⁵, il VdB non ci dice quale dei significati di, ad esempio, *giungere* è fondamentale, non lo dice il DIB, né il DAIC: bisogna aspettare il GRADIT, e il dizionario in un volume della Paravia, per vederlo scritto a caratteri rossi prima della definizione. Cer-

²³ Si assiste di anno in anno alla «sconfessione» della marca CO per molte accezioni da parte degli studenti universitari intervistati durante i corsi: la «buona scolarizzazione di livello secondario superiore», il «livello mediosuperiore di istruzione» è concetto mutevole.

²⁴ L'indice di dispersione è la misura della presenza del lemma nei diversi sottosettori dell'intero corpus su cui si fanno i rilevamenti statistici.

²⁵ *Prime parole. Dizionario illustrato di base della lingua italiana* compilato da De Mauro con due insegnanti, Elio D'Aniello e Gisella Moroni [1997], è il VdB definito con parole e 800 illustrazioni per i più giovani.

to chi conosce bene la lingua e ha una sensibilità spiccata per l'uso non doveva aspettare i dizionari di De Mauro per dire che *giungere* «arrivare» è significato fondamentale rispetto a *giungere* «congiungere» marcato CO, ma chi non ha sensibilità trova solo nei dizionari di De Mauro questa guida, perché per ora sono le uniche opere lessicografiche ad attribuire una marca di uso ad ogni accezione della voce.

Nel sito della rivista «Internazionale», da cui si può scaricare il NVdB, si può consultare anche il *Nuovo De Mauro*, un dizionario che, come si ricava dalle *Avvertenze* on line, prende le mosse dal *Dizionario della lingua italiana* «preparato e pubblicato, nel 2001, presso l'editore Paravia (poi Paravia-Bruno Mondadori, ora Mondadori-Pearson), un'opera in un unico volume, per proporre in sintesi l'essenziale dell'opera maggiore», cioè del GRADIT, passando da 260.000 lemmi semplici e 130.000 polirematiche, a 130.000 lemmi e 30.000 polirematiche.

Per allenare gli studenti ad esercitare la propria coscienza dell'uso di una parola si può proporre loro di controllare, ad esempio, se le parole AU nel NVdB sono AU anche nel *Nuovo De Mauro*. Dopotutto sono passati vent'anni tra il 1999 data dell'attribuzione delle marche d'uso nel GRADIT (e quindi nel *Dizionario della lingua italiana* e nel *Nuovo De Mauro*) e 50 dal primo VdB: le parole e le accezioni subiscono un cambiamento di fascia, verso il maggiore o minore uso. Ad esempio *adulto* che nel GRADIT 2007 è AU, nel NVdB e nel *Nuovo De Mauro* (2016) è FO; *compenso* che nel VdB era FO ed è FO nel GRADIT 2007 nell'accezione «corrispettivo in denaro per un lavoro svolto» è diventato AU nel NVdB e nel *Nuovo De Mauro*; *partigiano* che era FO nel NdB è ora AU e CO in tutte le accezioni più diffuse²⁶ *tronco* aggettivo nel senso di «troncato, mozzo, monco» e in quello figurato di «interrotto, lasciato a metà» era sceso da FO nel VdB a CO già nel GRADIT e sorprende molti

²⁶ E BU, basso uso nell'accezione «tifoso» e «AU TS stor. con particolare riferimento agli eventi della seconda guerra mondiale, chi partecipò alla resistenza contro il nazifascismo» nel *Nuovo De Mauro* mentre nel GRADIT 2007 la stessa accezione era FO TS.

giovani universitari del fatto che sia dato come CO e non BU (basso uso) nel NVdB e nel *Nuovo De Mauro*²⁷. È lo stesso De Mauro che, in un certo senso, ci esorta ai confronti nelle *Avvertenze*:

Il dizionario che si offre alla libera consultazione verrà progressivamente migliorato introducendovi correzioni e completamenti. A questo provvederà la redazione ma lo farà il più possibile raccogliendo osservazioni, suggerimenti, proposte delle persone che consultano il dizionario. La collaborazione di lettrici e lettori sarà preziosa per rifinire e perfezionare le singole voci e l'intera opera.

A parte questo aspetto da Wikizionario²⁸ – in realtà mai sviluppato –, il fatto che siano in rete, virtualmente vicini, sia il NVdB sia il *Nuovo De Mauro* è una preziosa occasione da sfruttare: l'insegnante ha l'immediata possibilità di giudicare se il significato in contesto di una parola fondamentale è quello FO oppure è CO e, soprattutto, lo può fare lo studente. La presenza piuttosto vistosa anche on line di polirematiche dovrebbe abituare lo studente a considerare le unità linguistiche formate da più parole dei lemmi, realtà frequente nei dizionari inglesi, più contenuta nei dizionari di lingue romanze.

Prendiamo la voce **giungere**: il corpo della voce è lo stesso nel GRADIT, nel *Dizionario* Paravia in un volume e nell'«Internazionale», ma le polirematiche – che nella versione digitale del dizionario in un volume non erano

²⁷ La marca CO, comune, è presa da molti in un senso più vicino ad AU, alto uso, di quanto non l'intenda De Mauro; solo quando si fa riflettere che un quarto, cioè 40.000 (47.060 nel GRADIT) delle accezioni dei lemmi singoli del *Nuovo De Mauro* hanno questa marca insieme a ben 30.000 polirematiche i parlanti comprendono che comune non vuol dire veramente comune. Quanto a BU, basso uso, nelle *Avvertenze* del *Nuovo De Mauro* si legge: «sono così marcati vocaboli rari, tuttavia circolanti ancora con qualche frequenza in testi e discorsi del Novecento».

²⁸ Come si legge nel sito <https://it.wiktionary.org/wiki/> è un dizionario nel quale i lettori possono essere anche autori. Bisogna rispettare alcune regole che caratterizzano il progetto; il Wikizionario italiano ha quasi raggiunto i 50.000 articoli e ha già superato i 100 file multimediali (informazione del 3 luglio 2017).

né definite né esemplificate²⁹ – sono ora on line corredate dalle definizioni e dagli esempi presenti nel GRADIT e nel dizionario monovolume a stampa; i «lavori in corso» si notano dal fatto che gli esempi delle polirematiche non sono in corsivo.

giungere

giùnl|gelre v.intr. e tr. av. 1250; dal lat. iūngĕre.

FO

1. v.intr. (*essere*) arrivare in un luogo: *giungere a casa, giungere alla meta* | sopraggiungere: *è giunto il momento, è giunta l'estate* | pervenire: *la notizia mi è giunta inaspettata; giungere al cuore, commuovere*

2. v.intr. (*essere*) spec. seguito da *a* e da un infinito, spingersi fino a un limite estremo e oltre: *sei giunto a tanto, giungere a compiere determinate azioni*

3. v.intr. (*essere*) OB LE avvenire, avverarsi

4. v.tr. CO congiungere, unire; *giungere le mani*: accostare le palme spec. in segno di devozione o preghiera

5. v.tr. LE aggiungere: *pur vai | giugnendo legne al foco* (Petrarca)

6a. v.tr. LE raggiungere: *sempre corsi, e mai non giunsi il fine* (Carducci)

6b. v.tr. LE colpire, ferire: *a due man prende | la crudel spada, e giunge il duca Arnolfo* (Ariosto) | fig., turbare: *io fui giunto | da' duo begli occhi* (Petrarca)

6c. v.tr. OB LE sorprendere

POLIREMATICHE

giungere a conoscenza

loc.v.

CO

venire a conoscenza: sono appena giunto a conoscenza della notizia

giungere all'orecchio

loc.v.

CO

apprendere, venire a conoscenza di qcs. per via indiretta: mi è giunta all'orecchio la notizia della tua partenza

²⁹ Mentre nella versione cartacea monovolume lo sono.

giungere in porto

loc.v.

CO

andare in porto

giungere nuovo

loc.v.

CO

di fatto o evento, essere, risultare inaspettato, inatteso: questa notizia mi giunge nuova.

Si può chiedere agli studenti perché la polirematica *a mani giunte* non compaia nella voce **giungere**, anche se contribuisce sicuramente alla marca d'uso CO di *giungere le mani*. La locuzione avverbiale, marcata CO, sta sotto **mano**, perché le polirematiche sono lemmatizzate e definite sotto il loro primo elemento pieno³⁰, e soltanto sotto di esso.

Come siano state isolate le polirematiche è detto in molti lavori fra cui Voghera [1994]. A De Mauro preme rilevare come il GRADIT abbia riunito le polirematiche in fondo alla voce per dar loro rilievo e far pulizia nella voce:

un rilevante apporto iniziale venne da Annibale Elia e dal suo gruppo di collaboratori salernitani. Impegnati da anni nello studio della lessicogrammatica, chiesi a loro di ricavare dai dizionari Trecani e dallo Zingarelli le polirematiche sparse un po' casualmente nel corpo delle definizioni dei vari lemmi, frammiste a esempi d'uso. Il progetto prevedeva che ciascuna polirematica sarebbe apparsa in coda a ognuno dei lemmi delle parole piene che la compongono, dunque *prendere fischi per fiaschi* deve apparire in calce a *prendere*, a *fischi* e a *fiaschi*. Il gruppo di Salerno contribuì con un primo nucleo di circa ventimila polirematiche, che poi nel lavoro redazionale sono state integrate e più che triplicate» [De Mauro 2017, 14].

È ben noto che nei dizionari digitali non è cruciale, dal punto di vista dell'utente, sotto quale lemma è registrata una locuzione formata da più parole piene, perché viene ritrovata sia cercando una delle parole sia cercando l'altra³¹ e viene

³⁰ Parole piene sono nomi, aggettivi, verbi, avverbi.

³¹ Si può estendere la ricerca di una parola nelle polirematiche nella

ritrovata comunque con la ricerca tuttotesto, mescolata a un po' di voci in cui compare, magari, in un esempio. Però dal punto di vista del metalessicografo linguista quale De Mauro era, la questione non è di poco conto, anzi è una di quelle che gli stavano a cuore e una di quelle su cui i continuatori della sua opera sono invitati a lavorare ancora.

Ci interessa però qui illustrare le ricerche avanzate possibili nel *Nuovo De Mauro* on line, che sono certo molto ridotte rispetto a quelle possibili sul GRADIT, ma generose rispetto alla media di quelle che è possibile fare sugli altri dizionari gratuiti on line³². Nelle *Avvertenze* la pagina dedicata alla *Ricerca avanzata* recita

All'interno del dizionario De Mauro Paravia (sic) è possibile effettuare ricerche avanzate utilizzando tre caratteri speciali: «?», «.» e «*».

? rappresenta un singolo carattere (compreso il carattere nullo)

. rappresenta un singolo carattere (escluso il carattere nullo)

* rappresenta zero o più caratteri

Al fine di non appesantire il lavoro del server il risultato della ricerca è limitato ai primi 500 risultati.

Per fare un esempio la ricerca ***ardo** produce una lista di 131 parole che vengono mostrare tutte: prima una serie in ordine alfabetico da *anacardo* a *traguardo*, poi una nuova serie da *azzardo* a *tardo* e infine 37 polirematiche elencate in un ordine alfabetico della prima parola piena e comprendenti fra gli altri *a bardosso*, *bietola cardonetto*, *riso sardonico*, dei già meno sorprendenti *alla gagliarda*, *nei riguardi di*, e dei prevedibili (digitalmente, non umanamente prevedibili salvo che per enigmisti provetti ed enciclopedici),

versione digitale del GRADIT 2007, fornita prima su Cd-ROM e poi su chiavetta USB. Il sistema di ricerca è fra i più interessanti e complessi del panorama italiano: le istruzioni per l'uso pubblicate a stampa raggiungono nel volume del 2003 90 pagine, ariose e piene di riproduzioni di videate. Nel GDU 2007 le *Funzionalità della versione digitale* stese da Clara Allasia per il manuale digitale sono solo 23 pagine, perché non ci sono più le videate.

³² Fra quelli monolingui di italiano il Sabatini-Coletti consente pure le ricerche con i caratteri speciali.

effetto di Ricardo, nardo sottile, ritardo mentale, ritmo lombardo.

Riflettendo si capisce che l'esempio fatto nella pagina delle *Avvertenze* e cioè «cercando ***orché**» si trovano tutti i lemmi che finiscono con **orché**» può fuorviare in quanto non ci sono altre parole italiane che contengano la sequenza di lettere *orché* e poi proseguano con altre lettere, mentre con la sequenza di lettere *ardo* questo può accadere perché «*rappresenta zero o più caratteri» e non è detto che debbano essere gli ultimi caratteri di una parola³³. La lista di polirematiche poi dimostra che la ricerca avviene sul lemma sotto cui sta la polirematica: avendo trovato *gagliardo* elenca la sua polirematica *alla gagliarda* e avendo individuato *riguardo* elenca poi le polirematiche *di riguardo, nei riguardi di, riguardo a*.

In una lingua ricca di morfologia derivativa come l'italiano i docenti dovrebbero incoraggiare le ricerche con i caratteri speciali, perché fanno fare scoperte giocose e istruttive. Si impara semiseriamente come è formata la parola italiana (o di un'altra lingua) e se una comune sequenza di lettere è in realtà un morfema con qualche funzione anche sul piano del significato semantico o grammaticale. Nei dizionari cartacei era possibile raggruppare le parole che iniziavano nello stesso modo e cercare fra queste i prefissi; le parole che finiscono in un certo modo si possono raggruppare solo con le versioni digitali³⁴.

4. *Proseguendo oltre il GRADIT*

Nel più volte citato articolo *Memorie del GRADIT* De Mauro dichiara:

Nelle sue introduzioni e postfazioni il GRADIT è riuscito in gran parte a onorare l'impegno di dichiarare le sue fonti e i

³³ Anche se questa è la situazione più comune nei dizionari digitali. Ad esempio il Sabatini-Coletti per la ricerca **ardo* elenca unicamente i 37 lemmi formati da una parola sola che terminano in *-ardo*.

³⁴ A meno di procurarsi un dizionario inverso a stampa.

criteri di trattamento, comprese le esclusioni, della immensa massa lessicale di una grande lingua di cultura: immensa non in senso vagamente enfatico ma, come il GRADIT aiuta a capire, oggettivamente non misurata e forse non misurabile» [De Mauro 2017, 17].

Ringrazia molto Sanguineti «lessicomane» e ricorda come il GRADIT è servito come base per la redazione del *Primo Tesoro della lingua letteraria italiana*, edito da Utet e Fondazione Bellonci nel 2007. Dagli scritti dei 60 vincitori del premio Strega da *Tempo di uccidere* (1947) di Flaiano a *Caos calmo* (2006) di Veronesi³⁵ Però i risultati dello spoglio hanno arricchito la versione 2007 del GRADIT di lemmi e prime attestazioni quali, fra gli altri, *extraterreno* (L. Romano, *Le parole tra noi leggere*, 1969), *gemmatura* (M. Bellonci, *Rinascimento privato*, 1985), *hélas* (E. Siciliano, *I bei momenti*, 1997), *imperlatura* (G. Montefoschi, *La casa del padre*, 1994), *magiarità* (C. Magris, *Danubio*, 1986), *schmarren* (N. Ginzburg, *Lessico familiare*, 1963).

Quella che può sembrare un'ossessione delle fonti è condizione imprescindibile per chi si è posto il compito di dare una marca d'uso per ogni accezione³⁶.

Nell'articolo del 2017 che ha assunto il sapore postumo di un testamento (meta)lessicografico, De Mauro quasi in chiusura menziona il «dare conto dell'ordinamento delle accezioni, che nell'opera di volta in volta in parte si fonda sulla diacronia, cioè sulla successione documentata delle accezioni, in parte è presentata come risultante di possibili estensioni o restrizioni e specificazioni delle accezioni» come una parte di progetto non realizzata³⁷. E prosegue:

³⁵ Più una selezione di altri 40 romanzi per portare il *Tesoro* a 100 opere.

³⁶ Anche se come osserva Rovere [2003, 5] queste fonti appaiono usate più per testimoniare la data di prima attestazione che sfruttate appieno – con un'ottica di linguistica dei corpora – per esemplificare le accezioni specialistiche e i loro intorni sintattici. L'esempio addotto da Rovere, il trattamento del verbo *investire* in psicologia e psicanalisi, è particolarmente efficace, dal momento che implica pure diverse costruzioni: investire qualcosa o qualcuno di qualcos'altro, investire qualcosa su qualcos'altro o qualcun altro.

³⁷ Il lavoro intorno al *Grande dizionario italiano dei sinonimi e con-*

Solo riuscendo a rintracciare la data di apparizione di ciascuna accezione e la consistenza della sua presenza in varie fasi storiche della lingua sarebbe possibile optare con decisione per una rappresentazione diacronica del succedersi delle accezioni: ma solo poche di queste e solo poche polirematiche è stato possibile datare con sicurezza [*ibidem*, 2017, 16-17].

È la strada che De Mauro indica a chi vuole proseguire il lavoro lessicografico come vuol fare anche l'Accademia della Crusca con il *Vocabolario dinamico dell'italiano moderno* (VoDIM)³⁸. Compito assolutamente non facile specie se si vuole ricorrere anche alla «storia documentabile dei dialetti che intridono l'uso dell'italiano»³⁹ per ordinare compiutamente le accezioni su base diacronica.

Riferimenti bibliografici

Baldelli, I. e Mazzetti, A.

1974 *Vocabolario minimo della lingua italiana per stranieri*, Firenze, Le Monnier.

Blumenthal, P. e Rovere, G.

1998 *PONS. Wörterbuch der italienischen Verben. Konstruktionen, Bedeutungen, Übersetzungen* (WIV), Stuttgart, Klett.

trari, con un'appendice di olonimi e meronimi (2010) deve avergli fatto maggiormente notare questa tensione fra microstrutture più diacroniche e altre più costruite su traslati, estensioni e restrizioni del significato.

³⁸ <http://www.accademiadellacrusca.it/it/eventi/crusca-torna-vocabolario-lessicografia-dinamica-dellitaliano-post-unitario>, consultato il 30 agosto 2017.

³⁹ De Mauro [2017, 17]. Alla presenza di parole di varietà regionali e derivanti da dialetti il GRADIT e, in misura leggermente minore, il dizionario in un volume del 2000 come pure il *De Mauro Internazionale*, dedicano attenzione. Si veda Marello e Sgroi [2015, 576-589] per un'analisi accurata delle marche RE regionale (usata 5.407 nel GRADIT) e quella DI dialettali (usata 338); delle 5 macroetichette diatopiche (sett., centro-sett., centr., centromerid., merid.) riferite a circa 900 macroregionalismi; delle 27 microetichette diatopiche, relative a 18 microaree regionali e a 9 microaree cittadine a livello sincronico, che riguardano circa 2.700 microregionalismi; più altre 10, che appaiono solo nell'ambito dell'«etimologia» e non tra le «marche d'uso», come *bolognese, bresciano, ferrarese, ladino, livornese, lucchese, modenese, romancio, torinese, valdostano*.

- 2016 *WIV Wörterbuch der italienischen Verben*, Nürnberg, Acolada, 2^a ed. (digitale) 2016; nuova ed. digitale 2017.
- Bortolini, U., Tagliavini, C. e Zampolli, A. (a cura di)
- 1972 *Lessico di frequenza della lingua italiana contemporanea* (LIF), Milano, Garzanti.
- Cardinale, U.
- 2002 *Recensione del Grande dizionario dell'uso* di T. De Mauro, in «Romance Philology», vol. 55, pp. 313-318.
- Chiari, I. e De Mauro, T.
- 2012 *The new basic vocabulary of Italian: problems and methods*, in «Rivista di statistica applicata / Italian Journal of Applied Statistics», vol. 22, 1, pp. 21-31.
- 2014 *The New Basic Vocabulary of Italian as a linguistic resource*, in R. Basili, A. Lenci e B. Magnini (a cura di), *First Italian Conference on Computational Linguistics CLiC-it 2014, 9-10 December 2014*, Pisa, Pisa University Press, pp. 113-116.
- D'Aniello, E., De Mauro, T. e Moroni, G.
- 1997 *Prime parole. Dizionario illustrato di base della lingua italiana*, Torino, Paravia.
- De Mauro, T.
- 1980 *Guida all'uso delle parole*, Roma, Editori Riuniti, 2^a ed. 1991.
- 1989 *I vocabolari ieri e oggi. I vocabolari oggi e domani*, in Direzione Ricerca Scientifica e Tecnologica Ibm Italia [1989, 7-50].
- 1997 *Dizionario avanzato dell'italiano corrente* (DAIC), Torino, Paravia.
- 1999 *Grande dizionario dell'Uso*, Torino, Utet.
- 2000 *Dizionario della lingua italiana*, Torino, Paravia.
- 2004 *Il dizionario di italiano compatto*, Torino, Paravia.
- 2005 *La fabbrica delle parole. Il lessico e problemi di lessicologia*, Torino, Utet.
- 2005 (a cura di), *Parole e numeri. Analisi quantitative dei fatti di lingua*, Roma, Aracne.
- 2010 *Grande dizionario italiano dei sinonimi e contrari*, con un'appendice di olonimi e meronimi, 2 voll., con CD-Rom, Torino, Utet.
- 2017 *Memorie del GRADIT*, in C. Marazzini (a cura di), *Utet: il laboratorio della parola*, Torino, Utet (stampa 2016), pp. 167-178; ristampato in «Lingua e stile», LII, giugno 2017, pp. 7-17.

- De Mauro, T. e Lo Cascio, V. (a cura di)
 1997 *Lessico e grammatica: teorie linguistiche e applicazioni lessicografiche*, Roma, Bulzoni.
- De Mauro, T., Mancini, F., Vedovelli, M. e Voghera, M.
 1993 *Lessico di frequenza dell'italiano parlato* (LIP), Milano, Etaslibri.
- De Mauro, T. e Moroni, G.G.
 1996 *DIB Dizionario di base della lingua italiana*. Torino, Paravia.
- De Palo, M.
 1997 *Il Vocabolario di Base a confronto con il Lessico Italiano del Parlato*, in De Mauro e Lo Cascio [1997, 395-411].
- Direzione Ricerca Scientifica e Tecnologica Ibm Italia
 1989 *Vocabolario Elettronico della Lingua Italiana* (VELI), Milano, Direzione Comunicazioni Ibm Italia.
- Gensini, S. e Vedovelli, M. (a cura di)
 1983 *Teoria e pratica del glotto-kit. Una carta d'identità per l'educazione linguistica*, Milano, FrancoAngeli, 3ª ed. 1991.
- Juilland, A. e Traversa, V.
 1973 *Frequency Dictionary of Italian Words*, The Hague, Mouton.
- Katerinov, K., Boriosi Katerinov, M.C. e Sciarone, A.G.
 1991 *Calcolo di rendimento di liste di base: italiano parlato, italiano scritto, livello soglia*, Perugia, Guerra.
- Marello, C. e Sgroi, S.C.
 2015 *La regionalità nella lessicografia italiana*, in M. Salvati e L. Sciolla (a cura di), *L'Italia e le sue regioni (1945-2011) Pratiche, memoria e varietà linguistica*, Roma, Istituto della Enciclopedia Italiana, pp. 569-589.
- Reiske, H.
 1974 *VHS-Italienisch Unterricht und VHS-Zertifikat Italienisch*, Frankfurt, Pädagogische Arbeitsstelle Deutscher Volkshochschul-Verband.
- Rovere, G.
 2003 *Das GWDS und der Grande Dizionario Italiano dell'Uso (GDU)*, in H.E. Wiegand (a cura di), *Untersuchungen zur kommerziellen Lexikographie der deutschen Gegenwartssprache I. «Duden. Das große Wörterbuch der deutschen Sprache in zehn Bänden»*. Print- und CD-ROM-Version, (Lexicographica. Series Maior, Band 113), Tübingen, Niemeyer, pp. 61-80.

- Sciarone, A.G.
 1977 *Vocabolario fondamentale della lingua italiana*, Bergamo, Minerva Italica; ed. 1995, Perugia, Guerra Edizioni.
- Vedovelli, M.
 1993 *Confronti fra il LIP e le altre liste di frequenza dell'Italiano*, in LIP, pp. 119-147.
- Voghera, M.
 1994 *Lessemi complessi: percorsi di lessicalizzazione a confronto*, in «Lingua e stile», XXIX, 2, pp. 185-214.
- Voghera, M., Iacobini, C., Savy, R., Cutugno, F., De Rosa, A. e Alfano, I.
 2014 *VoLIP: A Searchable Corpus of Spoken Italian*, in L. Veselovská e M. Janebová (a cura di), *Complex Visibles Out There. Proceedings of the Olomouc Linguistics Colloquium 2014: Language Use and Linguistic Structure*, Olomouc, Palacký University, pp. 627-639.